Spectrum-Aware Mobility Management in Cognitive Radio Cellular Networks

Won-Yeol Lee, *Student Member*, *IEEE*, and Ian F. Akyildiz, *Fellow*, *IEEE*

Abstract—Cognitive radio (CR) networks have been proposed as a solution to both spectrum inefficiency and spectrum scarcity problems. However, they face several challenges based on the fluctuating nature of the available spectrum, making it more difficult to support seamless communications, especially in CR cellular networks. In this paper, a spectrum-aware mobility management scheme is proposed for CR cellular networks. First, a novel network architecture is introduced to mitigate heterogeneous spectrum availability. Based on this architecture, a unified mobility management framework is developed to support diverse mobility events in CR networks, which consists of spectrum mobility management, user mobility management, and intercell resource allocation. The spectrum mobility management scheme determines a target cell and spectrum band for CR users adaptively dependent on time-varying spectrum opportunities, leading to increase in cell capacity. In the user mobility management scheme, a mobile user selects a proper handoff mechanism so as to minimize a switching latency at the cell boundary by considering spatially heterogeneous spectrum availability. Intercell resource allocation helps to improve the performance of both mobility management schemes by efficiently sharing spectrum resources with multiple cells. Simulation results show that the proposed method can achieve better performance than conventional handoff schemes in terms of both cell capacity as well as mobility support in communications.

Index Terms—Cognitive radio, spectrum pool, handoff, intercell resource allocation, spectrum mobility management, user mobility management.

1 Introduction

Wireless spectrum is currently regulated by governmental agencies and is assigned to license holders or services on a long-term basis over vast geographical regions. Recent research has shown that a large portion of the assigned spectrum is used sporadically leading to under utilization and wastage of valuable frequency resources [1]. To address this critical problem, the Federal Communications Commission (FCC) has recently approved the use of unlicensed devices in licensed bands [14]. This new area of research foresees the development of cognitive radio (CR) networks to further improve spectrum efficiency. The basic idea of CR networks is that the unlicensed devices (also called CR users) share the licensed spectrum without interfering with the transmission of other licensed users (also known as primary users) [2]. If this band is found to be occupied by a licensed user, the CR user moves to another spectrum hole to avoid interference, which is referred to as *spectrum mobility*.

This concept has been widely investigated to solve the exponential data traffic growth in the current cellular network [4], [5]. The main difference between classical and CR cellular networks lies in spectrum mobility, which gives rise to a new type of handoff in CR cellular networks, the so-called *spectrum handoff*. In [6], a proactive spectrum handoff approach is proposed where CR users predict future spectrum availability based on the past channel histories, and intelligently switch the channel prior to the appearance of primary users, which minimizes disruptions to primary users and maintains reliable communication at CR users. A spectrum handoff scheme based on a discrete-time Markov chain is developed to reduce a forced termination probability of CR transmission [7]. IEEE 802.22, the first CR standard for CR networks, maintains a backup channel list so as to provide the highest probability of finding an available spectrum band within the shortest time [3]. In [8], an algorithm for updating the backup channel list is developed to find the idle spectrum bands fast and reliably by cooperating with neighbor CR users. However, recent research mentioned above has mainly focused on spectrum mobility, but does not consider the effect of mobile users across multiple cells.

In the cellular network, mobility management, especially a handoff scheme is one of the most important functions. Thus, much research on cellular networks has explored the handoff issues, mainly focusing on cell selection and resource management in the last couple of decades [9]. Although diverse cell selection methods have been proposed to support seamless handoff schemes while maximizing the network capacity [10], [11], [12], [13], all of them are based on the classical multicell-based networks and do not consider the fluctuating nature of spectrum resource in CR networks. Especially, no special attention is given to either time and location-varying spectrum availability or switching delay in traversing the spectrum distributed over a wide frequency range, which makes conventional handoff schemes infeasible in CR cellular networks.

As a result, spectrum mobility and user mobility must be jointly considered in designing a mobility management scheme for CR cellular networks, which constitutes an

- *The authors are with the Broadband and Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, 777 Atlantic Dr., NW, Atlanta, GA 30332-0250. E-mail: {wylee, ian}@ece.gatech.edu.*

Manuscript received 26 May 2009; revised 25 Oct. 2010; accepted 10 Jan. 2011; published online 17 Mar. 2011.
For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number TMC-2009-05-0193.
Digital Object Identifier no. 10.1109/TMC.2011.69.

1536-1233/12/$31.00 © 2012 IEEE Published by the IEEE CS, CASS, ComSoc, IES, & SPS

important but unexplored topic in CR networks to date. There are the following challenges:

- *Heterogeneous mobility events.* CR networks are required to provide two different handoff types: classical intercell handoff due to physical user mobility and spectrum handoff due to spectrum mobility. Thus, it necessitates a unified mobility management framework to exploit two handoff types adaptively to mobility events.
- *Dynamic spectrum availability.* Spectrum availability in CR networks varies over time and space, making it more difficult to provide seamless and reliable communications to mobile users traversing across multiple cells. For efficient mobility management, CR networks need to mitigate this dynamic spectrum availability by performing mobility management adaptively dependent on the heterogeneous network conditions.
- *Broad range of available spectrum.* In CR networks, available spectrum bands are not contiguous and found over a wide frequency range. Thus, when CR users switch their spectrum bands, they need to reconfigure the operating frequency of the radio frequency (RF) front-end so as to tune to a new spectrum band, leading to a switching delay much longer than that in classical wireless networks.

To address the challenges mentioned above, we propose a spectrum-aware mobility management scheme for CR cellular networks. First, we propose a novel CR cellular network architecture based on the spectrum pooling concept, which mitigates the heterogeneous spectrum availability. Based on this architecture, a unified mobility management framework is defined so as to support diverse mobility events in CR networks, consisting of intercell resource allocation, spectrum mobility management, and user mobility management functions. Through intercell resource allocation, each cell determines its spectrum configuration to improve mobility as well as total capacity. To support spectrum mobility while maintaining maximum cell capacity, the spectrum mobility management function determines a proper handoff type and target cell for CR users experiencing PU activities by considering both spectrum utilization and stochastic connectivity models. On the other hand, user mobility management mainly focuses on spectrum heterogeneity in space, and offers a switching cost-based handoff decision mechanism to minimize service quality degradation in mobile users.

The rest of the paper is organized as follows: Sections 2 and 3 present the proposed network architecture and mobility management framework for CR networks, respectively. Handoff models in the proposed framework are presented in Section 4. In Sections 5 and 6, novel spectrum and user mobility management methods are proposed, respectively. Performance evaluation and simulation results are presented in Section 7. Finally, conclusions are presented in Section 8.

2 Proposed Network Architecture

The cellular network is the most successful wireless technology, but currently suffers from increasing data traffic as data hungry smartphones proliferate. The CR technology is considered as a promising solution to this data explosion problem in the current cellular network.

The CR cellular network is supposed to be deployed in several ways. First, it can be applied to the unused TV spectrum bands, the so-called *TV white spaces*, as the FCC recently allowed unlicensed devices to use them [5], [14]. Second, while the ultrabroadband cellular technology such as 3rd Generation Partnership Project (3GPP) Long-Term Evolution-Advanced (LTE-Advanced) requires up to 100 MHz per channel [15], the amount of wideband spectrum is limited. The CR technology enables bandwidth aggregation by sharing spectrum owned by other cellular operators, or opportunistically utilizing unused spectrum bands licensed to other services such as digital TV, and public safety [16]. Finally, in the current cellular networks, the base station (BS) has only an RF unit, and a digital unit for all communication functionalities is implemented in a separate central server [17]. As a result, the cost of BSs will be cheap enough for anybody to install anywhere. This allows a new type of a mobile virtual operator based on CR, which operates its own BSs in a local area without spectrum licenses.

In this paper, we mainly focus on the mobility issues in CR cellular networks. The system model used in this paper is described in the following sections.

2.1 Basic System Model

In this paper, we consider infrastructure-based CR networks consisting of multiple cells. Each cell has a single BS and its CR users. In this architecture, CR users observe their radio environments and report the results to the BS. Accordingly, the BS determines proper actions in support of a upper-level control node, such as the mobility management entity (MME) in 3GPP LTE [15]. CR users have a single wideband RF transceiver that can sense multiple contiguous spectrum bands at the same time without RF reconfiguration. Each CR user m needs K_m channels to satisfy its QoS requirement.

All spectrum bands are assumed to be licensed to different primary networks. Furthermore, each spectrum can have multiple primary networks that are operated independently in the different region, called *PU activity region*. For example, the cell coverage of each BS is considered as the PU activity region. Since most of primary networks such as cellular networks or TV broadcasting networks have a fixed service coverage, the PU activity region is generally assumed to be fixed [18].

Generally, each PU activity region in the licensed band has ON and OFF states where ON (Busy) state represents the period used by primary users and an OFF (Idle) state represents the unused period. In this paper, we assume that the length of ON and OFF periods at PU activity region k in spectrum band j is exponentially distributed with means $1/\alpha(j,k)$ and $1/\beta(j,k)$, respectively [19], [20], [21]. Then, the idle probability in the licensed band $P^{\text{off}}(j,k)$ can be expressed as $\alpha(j,k)/(\alpha(j,k)+\beta(j,k))$. Since the state transitions are detected by periodic sensing operations in every interval of length Δt, the PU activity in the licensed band is modeled as a two-state Markov chain.

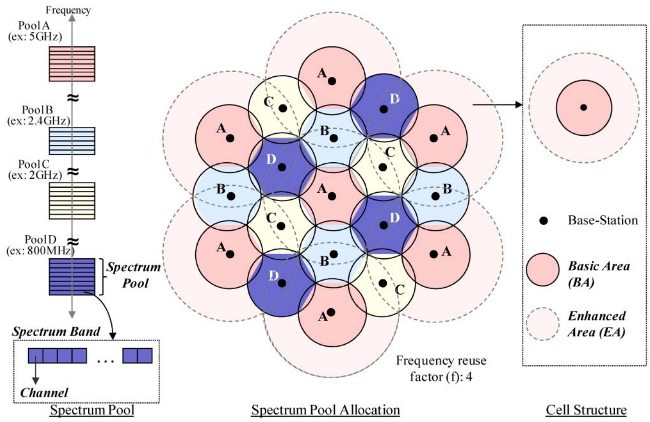


Fig. 1. Spectrum pool-based CR network architecture.

2.2 Spectrum Pool Structure

In this section, we present the proposed spectrum pool and cell architectures and a corresponding capacity model.

2.2.1 Spectrum Pool Architecture for Spectrum Mobility

To prevent intercell interference, classical cellular networks generally adopt an interference coordination scheme where each cell uses different spectrum band with its neighbor cells [15]. This concept can be also applied to CR cellular networks. If the spectrum bands are contiguous and located in a relatively narrow frequency range, mobile users can switch the spectrum without changing their RF front-ends. However, since available spectrum bands in CR networks are spread over a wide frequency range as explained in Section 1, CR users need to reconfigure the operating frequency of their RF front-end whenever they detect PU activities in the current band, resulting in a significant switching latency.

To solve the problem in spectrum mobility, we modify a conventional spectrum pooling concept, known as the most suitable structure to adapt to the dynamic radio environment in CR networks [22], [23], for handling both spectrum and user mobilities in a multicell environment. The main components of the spectrum pool are defined as follows (Fig. 1):

- *Spectrum pool:* A set of contiguous licensed *spectrum bands*, each of which consists of multiple *channels*.
- *Spectrum band:* A basic bandwidth unit for operating a certain wireless access technology such as 5 MHz WCDMA band operating at 2.1 GHz.
- *Channel:* A minimum logical unit of wireless resource that mobile users can access through multiple access schemes. Each channel has the identical capacity.

2.2.2 Cell Architecture for User Mobility

In the proposed architecture, spectrum pools are assigned to each cell exclusively with its neighbor cells with a predetermined frequency reuse factor, f, as shown in Fig. 1. Although this architecture supports a seamless transition between spectrum bands within the pool, we still has difficulty in providing seamless communication to CR users moving across different cells. To address this problem, we introduce two different types of cell coverage as depicted in Fig. 1:

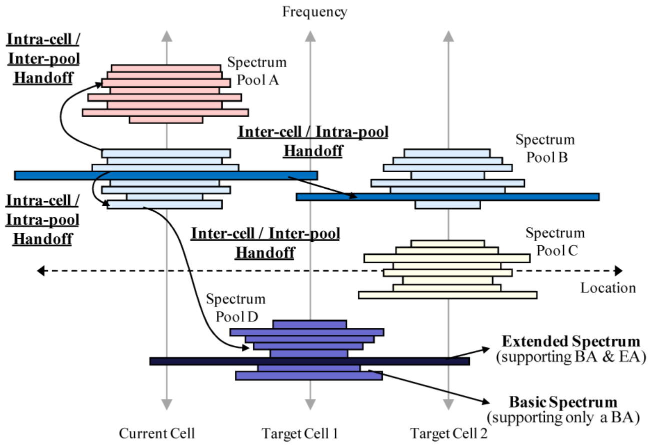


Fig. 2. Different handoff types in CR networks.

- *Basic area (BA):* A basic coverage not overlapped with that of neighbor cells.
- *Extended area (EA):* A larger coverage overlapped with the basic areas of neighbor cells.

Based on this architecture, each cell has multiple *basic spectrum* bands accessible only within BA and at least one *extended spectrum* band supporting both BA and EA, as shown in Fig. 2. The EA of the current cell is adjacent to those of other cells having the same spectrum pool, referred to as *extended neighbors*. The use of EA helps to improve the mobility performance significantly by maintaining the operating frequency of mobile users.

2.2.3 Spectrum and Cell Capacities

The basic spectrum offers $N_i^{\max}(j)$ channels in BA. To support the same number of channels in the larger coverage, the extended spectrum should use a higher transmission power than the basic spectrum. Assume that the extended spectrum j at spectrum pool i supports $N_i^{\max}(j)$ channels for the users in EA. Then it can support more channels, $\rho N_i^{\max}(j)$, to the users in BA due to the shorter distant from the BS where ρ is greater than unity and is determined dependent on the transmission power and the minimum signal strength for decoding.

The *cell capacity* is defined as the number of currently all available channels in the cell, i.e., the sum of channels in all spectrum bands currently not occupied by the primary networks in the cell. The use of the extended spectrum helps to improve the cell capacity. However, since the extended neighbors are located within the interference range of the extended spectrum in the current cell, the extended spectrum of the current cell cannot be used in its extended neighbors so as to avoid intercell interference.

2.3 Handoff Types

Mobility management in classical cellular networks is closely related to user mobility. However, CR networks have another unique mobility event, the so-called *spectrum mobility*. By taking into account both mobility events based on the proposed network architecture, we define four different types of handoff schemes as shown in Fig. 2:

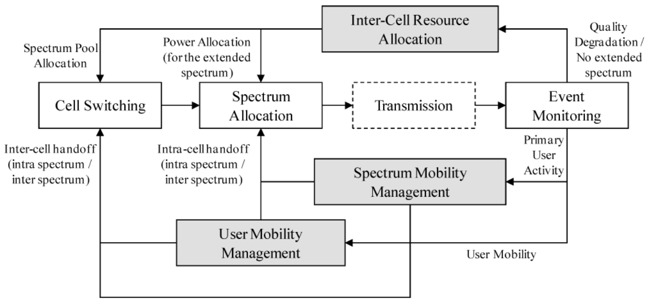


Fig. 3. The proposed mobility management framework.

- *Intracell/intrapool handoff.* The CR user moves to a spectrum band in the same spectrum pool without switching a serving BS.
- *Intercell/intrapool handoff.* The CR user switches its serving BS to a neighbor BS without changing the spectrum pool.
- *Intercell/interpool handoff.* The CR user switches its serving BS to a neighbor BS, which has a different spectrum pool.
- *Intracell/interpool handoff.* The CR user changes its spectrum bands from one spectrum pool to another within the current cell.

Each handoff type is related to different mobility event, and its performance is mainly dependent on both network and user conditions, such as resource availability, network capacity, user location, etc. Thus, CR networks require a unified mobility management scheme to exploit different handoff types adaptively to the dynamic nature of underlying spectrum bands, which will be explained in Section 3.

3 Mobility Management Framework

3.1 Overview

Compared to the classical cellular network, the CR network requires more complicated mobility management functionalities due to the dynamic spectrum environment and heterogeneous handoff types, as shown in Fig. 3. These functionalities are initiated by three different events: user mobility, spectrum mobility, and quality degradation. Here, *user mobility* is defined as the event that a mobile CR user transfers an ongoing connection from one BS to another as it approaches the cell boundary. On the contrary, *spectrum mobility* is referred to as the event that CR users switch their spectrum due to the PU activity. Each BS detects one of these events by monitoring current spectrum availability and the quality variation of current transmissions and performs a proper mobility management function accordingly.

In the case of user and spectrum mobility events, CR networks decide on a proper handoff type for their mobile users by performing user and spectrum mobility management functions, respectively. According to the decision, CR users need either to select a target cell (*cell selection*) or to determine the best available spectrum (*spectrum allocation*).

If a current cell does not have enough spectrum resource due to either PU activity or increase in users, the BS performs an *intercell resource allocation* through coordination with its neighbor cells. Through this operation, the cell can obtain the additional spectrum pool. This concept has been widely studied in [24], [25], and [26]. To mainly focus on the performance of mobility management, we assume that each cell has a single spectrum pool.

Besides the above spectrum sharing capability, intercell resource allocation necessitates a unique functionality to maintain the hierarchical cell structure, which is explained in the following section.

3.2 Intercell Resource Allocation

Since each cell has time-varying wireless resource because of the dynamic nature of underlying spectrum in CR networks, it cannot have a permanent extended spectrum. Furthermore, as explained in Section 2, the extended spectrum of the current cell cannot be used in its extended neighbors, leading to decrease in their capacity. As a result, CR networks necessitate an intercell resource allocation scheme to select and maintain the extended spectrum. Although global optimization in every spectrum change achieves optimal allocation, it requires a huge computational complexity and also causes a high communication overhead due to frequent spectrum switching. Instead, we consider the stochastic characteristics of spatial and temporal spectrum availabilities, and develop a distributed intercell resource allocation method, which improves total network capacity as well as mobility support, i.e., the availability of the extended spectrum. The following are the procedures of the proposed method:

1. Initially, all available spectrum bands in current cell i are considered as basic spectrum bands.
2. CR users can access the spectrum band only when all PU activity regions in the BA of the current cell are idle. Thus, the expected capacity of spectrum band j at cell i is defined as follows:

$$C_i(j) = N_i^{\max}(j) \cdot \prod_{k \in \mathcal{A}_i^{\mathrm{B}}(j)} P^{\mathrm{off}}(j,k), \quad (1)$$

where $\mathcal{A}_i^{\mathrm{B}}(j)$ is a set of the PU activity regions of spectrum j in the BA of cell i.

3. Once the extended spectrum is lost due to the PU activity, intercell spectrum sharing is performed to find a new spectrum, which takes time due to information exchange with its neighbor cells. Thus, the reliability of the extended spectrum can be expressed as the ratio of an average idle time in the extended spectrum band to total time including an intercell spectrum sharing delay as follows:

$$R_i(j) = \frac{\frac{1}{\sum_{k \in \mathcal{A}_i^{\mathrm{E}}(j)} \beta(j,k)}}{\frac{1}{\sum_{k \in \mathcal{A}_i^{\mathrm{E}}(j)} \beta(j,k)} + T^{\mathrm{inter}}}, \quad (2)$$

where $\mathcal{A}_i^{\mathrm{E}}(j)$ is a set of the PU activity regions of spectrum j in the EA of cell i, $1/\sum_{k \in \mathcal{A}_i^{\mathrm{E}}(j)} \beta(j,k)$ represents the average idle period of the extended spectrum j at cell i, and T^{inter} is the intercell resource allocation delay.

TABLE 1
Symbols Used for the Analytical Modeling

Symbols	Descriptions
N_i^{b}	Total number of channels used in the BA of the cell i
N_i^{e}	Total number of channels used in the EA of the cell i
$N_i^{\max}(j)$	Maximum number of channels in the spectrum band j at the BA of the cell i
$\alpha(j,k)$	PU activity (busy $\to$ idle) at the area k of the spectrum band j
$\beta(j,k)$	PU activity (idle $\to$ busy) at the area k of the spectrum band j
ρ	Channel gain of users in BA at the extended spectrum
Δt	Sensing interval (sensing operation in every Δt)

4. Each cell prefers an extended spectrum with higher reliability. However, once the current cell determines the extended spectrum, its extended neighbors cannot use that spectrum, and hence lose their capacity. To describe these features, we develop a novel metric for the expected gain, which can be expressed as the product of the spectrum reliability of the extended spectrum in the current cell and a ratio of the capacity gain in the current cell to the sum of capacity loss in extended neighbors as follows:

$$G_i(j) = R_i(j) \cdot \frac{\rho N_i^{\max}(j) \prod_{k\in\mathcal{A}_i^{\mathrm{E}}(j)} P^{\mathrm{off}}(j,k) - C_i(j)}{\sum_{i'\in\mathcal{N}_i^E} [C_{i'}(j) \cdot N_{i'}^{\max}(j)]}, \quad (3)$$

where $\mathcal{N}_i^E$ is a set of the extended neighbors of cell i.

5. The current cell considers the expected capacity gain over all available spectrum bands and chooses the extended spectrum j^* to satisfy the following condition:

$$j^* = \arg\max_{j\in\mathcal{S}_i} G_i(j), \quad (4)$$

where $\mathcal{S}_i$ is a set of currently idle spectrum at cell i.

In the following sections, we introduce handoff models according to the switching latency and then propose spectrum and user mobility management schemes. For ease of presentation, the important symbols used in the subsequent discussion are summarized in Table 1.

4 Spectrum Handoff Modeling

According to the mobility events, each handoff scheme necessitates different strategies as follows:

- *Proactive handoff.* When CR users detect handoff events, they perform handoff procedures while maintaining communications. After CR users make all decisions on the handoff, they cut off communication channels and switch to a new spectrum band or a new BS. User mobility and cell overload are the examples of proactive handoff events. Most of classical handoff schemes are based on the proactive approach.
- *Reactive handoff.* CR users should stop the transmission immediately in the reactive handoff event. Then, they make decisions and perform the handoff. As a result, this handoff has an additional handoff delay, unlike the proactive approach. In the case that the primary user appears in the spectrum, the CR network should initiate the reactive handoff by immediately vacating the spectrum to avoid interference, and then performing decision on a new available band.

Based on these strategies, the handoff schemes defined in Section 2 can be modeled as follows:

4.1 Intracell/Intrapool Handoff

Intracell/intrapool handoff occurs when primary users are detected in the spectrum. Thus, it is implemented in a reactive approach. First, this handoff approach requires a preparation time to determine the handoff type (d_{prep}). After that, for sensing operations, CR users need to wait for the next sensing cycle, called a sensing synchronization time ($d_{\mathrm{syn}}^{\mathrm{sen}}$). Then, they sense the spectrum bands in the pool (d_{sen}) and determine the proper spectrum (d_{dec}). Finally, CR users move to a new spectrum band and resume transmission after the synchronization to the transmission schedule on that spectrum ($d_{\mathrm{syn}}^{\mathrm{tx}}$). Since spectrum bands in the pool are contiguous, CR users can switch the spectrum without reconfiguring their RF front-ends, and hence the physical spectrum switching delay is negligible. In summary, the latency for intracell/intrapool handoff (Type 1) can be expressed as follows:

$$D_1 = d_{\mathrm{prep}} + d_{\mathrm{syn}}^{\mathrm{sen}} + d_{\mathrm{sen}} + d_{\mathrm{dec}} + d_{\mathrm{syn}}^{\mathrm{tx}}. \quad (5)$$

4.2 Intracell/Interpool Handoff

If CR BSs can exploit multiple spectrum pools, intracell/interpool handoff may happen in PU activity. If the current spectrum pool does not have enough spectrum resource due to PU activity, CR users detecting PU activities switch to another spectrum pool in the current cell. This is also a reactive handoff. Thus, its handoff latency is similar to that of the intracell/intrapool handoff as follows (Type 2):

$$D_2 = d_{\mathrm{prep}} + d_{\mathrm{recfg}} + d_{\mathrm{syn}}^{\mathrm{sen}} + d_{\mathrm{sen}} + d_{\mathrm{dec}} + d_{\mathrm{syn}}^{\mathrm{tx}}. \quad (6)$$

However, unlike the intracell/intrapool handoff, this scheme requires the reconfiguration of RF front end since each spectrum pool is placed in the different frequency range. Usually, reconfiguration takes longer time than other delay components.

4.3 Intercell/Interpool Handoff

This handoff scheme is similar to that in classical cellular networks, which is required for CR users moving across multiple cells. To determine a target cell, a mobile CR user needs to observe the signals from neighbor cells during its transmission. However, since neighbor cells use different spectrum pools, the mobile CR user should stop its transmission and reconfigure its RF front-end in every observation of neighbor cells, which is a tremendous overhead in handoff. Thus, instead of this mobile station-controlled method, a network-controlled approach is more feasible for intercell/interpool handoff, where the BS determines the target cell based on the stochastic user information, which is explained in Section 5. Consequently, mobile CR users need a single reconfiguration time. In this case, the BS prepares the handoff in advance according to user mobility. Thus, this is a proactive handoff and does not require the handoff preparation time d_{prep} used in previous reactive handoff types as follows (Type 3):

$$D_3 = d_{\mathrm{recfg}} + d_{\mathrm{syn}}^{\mathrm{sen}} + d_{\mathrm{sen}} + d_{\mathrm{dec}} + d_{\mathrm{sych}}^{\mathrm{tx}}. \quad (7)$$

Furthermore, PU activities can initiate this handoff scheme in special reactive events. First, when all spectrum pools in the current cell are overloaded due to PU activity, the BS forces CR users to move to neighbor cells. This is exactly the same procedure as the intracell/interpool handoff, and requires D_2 handoff latency. Second, if a PU activity is detected in the extended spectrum, CR users in the extended spectrum do not have other available band in the current cell, and hence should switch to the neighbor cells. In this case, they lose a control channel as well. To solve this problem, the BS determines handoff information and sends it to a selected target cell. Then, the target cell broadcasts the advertisement message for the CR user through its control channel. In this scenario, CR users need one or more reconfigurations of the RF front end until it hears the advertisement message. Also in every reconfiguration, CR users monitor the control channel for a certain time (d_{lis}). The latency in this case (Type 4) can be expressed as follows:

$$D_4 = d_{\mathrm{prep}} + \gamma(d_{\mathrm{recfg}} + d_{\mathrm{lis}}) + d_{\mathrm{syn}}^{\mathrm{sen}} + d_{\mathrm{sen}} + d_{\mathrm{dec}} + d_{\mathrm{syn}}^{\mathrm{tx}}. \quad (8)$$

Due to multiple reconfigurations, intercell/interpool handoff in this case shows the worst performance in terms of switching latency. γ is dependent on the searching order of neighbor cells. In this paper, the order is randomly chosen, and hence γ is considered as $(f+1)/2$ on average where f is a frequency reuse factor.

4.4 Intercell/Intrapool Handoff

This handoff happens when mobile CR users in EA successfully switch to the extended neighbors. This is also a proactive handoff. Furthermore, a new target cell is an extended neighbor that uses the same spectrum pool as the current cell, and hence reconfiguration is not required. Therefore, the latency for intercell/intrapool handoff scheme (Type 5) can be expressed as follows:

$$D_5 = d_{\mathrm{syn}}^{\mathrm{sen}} + d_{\mathrm{sen}} + d_{\mathrm{dec}} + d_{\mathrm{syn}}^{\mathrm{tx}}. \quad (9)$$

In this handoff, the latency is significantly reduced compared to that in other cases. Thus, this type of handoff is more advantageous to mobile CR users, and hence improves mobility in CR networks.

5 Spectrum Mobility Management in Cognitive Radio Networks

5.1 Overview

Spectrum mobility is the unique characteristic in CR networks. When primary users appear in the spectrum, CR users generally change its spectrum band without switching the BS. However, since CR networks have time-varying spectrum availability, each cell may not have enough spectrum band to support current users. To solve this problem, an admission control scheme is proposed in [27]. However, in the proposed architecture, CR users can have another option, cell switching, because of the hierarchical spectrum structure described in Section 2.2. Here, we propose a spectrum mobility framework by considering both spectrum and cell switching methods.

When the PU activity is detected in the cell, the BS needs to check if it has enough spectrum resource for intracell/intrapool handoff. If the cell has enough spectrum resource, the BS performs the intracell/intrapool handoff for all users requiring new spectrum bands (Type 1). Otherwise, some of current users are forced to move to the neighbor cells. If the PU activity is detected in the extended spectrum, all users in EA need to perform intercell/interpool handoff (Type 2) regardless of current spectrum resource since they cannot find other available spectrum bands for switching in that area.

After the user selection, the selected users need to find the proper target cell. Unlike the classical handoff, CR users cannot observe the signal strength from other neighbor cells while maintaining the connection to current cells. Instead, CR users select the new BS based on the stochastic connectivity model, which is explained in Section 5.3.

The intracell/intrapool handoff is exactly same as the spectrum decision proposed in [27], and hence out of scope in this paper. In the following sections, we will describe the user and cell selections for the intercell/interpool handoff scheme in more detail.

5.2 User Selection

Let $\mathcal{S}_i$ be a set of the currently available spectrum in the cell i. Then, the number of unused channels in the available spectrum bands at the cell i, N_i^{av}, can be expressed as $\sum_{j\in\mathcal{S}_i} N_i^{\max}(j) - (N_i^{\mathrm{b}} + \rho N_i^{\mathrm{e}})$. Here, N_i^{b} and N_i^{e} are the numbers of channels used in the BA and EA of cell i, respectively, and can be obtained as follows:

$$N_i^{\mathrm{b}} = \sum_{m\in\mathcal{U}_i^{\mathrm{b}}} K_m, \quad N_i^{\mathrm{e}} = \sum_{m\in\mathcal{U}_i^{\mathrm{e}}} K_m, \quad (10)$$

where $\mathcal{U}_i^{\mathrm{b}}$ and $\mathcal{U}_i^{\mathrm{e}}$ are the sets of users in BA and EA, respectively.

If the number of new channels required for spectrum switching N^{req} is less than N_i^{av}, i.e., the cell is not overloaded, CR users just perform the intracell intrapool handoff. As explained in Section 5.1, the users in EA should move to the neighbor cells when they detect the PU activity, and hence their channels are not counted in N^{req}.

When detecting cell overload, the current cell forces some of its users to be out to neighbor cells. In this case, to minimize the number of users to be switched, the overloaded cell chooses users occupying more spectrum resources until it becomes free from cell overload as follows:

- If $N_i^{\mathrm{av}} < N^{\mathrm{req}} \le N_i^{\mathrm{av}} + \rho N_j^{\mathrm{e}}$, CR users using $\lceil \frac{N^{\mathrm{req}} - N_i^{\mathrm{av}}}{\rho} \rceil$ channels in EA need to be selected and moved to the neighbor cells. As the users stay in EA for a longer time, cell capacity becomes lower. Also, these users have a higher probability to be interrupted by the PU activity. Furthermore, the users having more channels reduce the number of users that the current cell can admit. Thus, the BS selects the users in EA with the longest expected staying time as well as the highest channel occupancy. As a result, a user selection metric can be obtained as $K_m \cdot d_m/v_m$ where d_m is the expected moving distance of user m to the cell boundary, which is dependent on the user mobility model. v_m is the velocity of user m. The BS chooses users in EA with the largest decision metric, repeatedly until it can avoid the cell overload state.

- If $N^{\rm req} > N_i^{\rm av} + \rho N_i^{\rm e}$, it is not enough to select all users in EA. To avoid dropping or blocking connections due to the cell overload, the BS hands over some of users in BA to its neighbor cells. Unlike the previous case, the BS selects CR users using $N^{\rm req} - (N_i^{\rm av} + \rho N_i^{\rm e})$ channels with the shortest expected staying time in BA since they are highly likely to move to EA, which will require more spectrum resource. Similar to the previous case, it is more advantageous to hand over the users with more channels. Thus, the BS chooses CR users in BA with the smallest decision metric, $d_m/(v_m \cdot K_m)$.

5.3 Cell Selection

One of main challenges in CR mobility management is how to determine a proper target cell. Since each spectrum pool is distributed over a wide frequency range, CR users need to reconfigure their RF front-ends for monitoring the signals from neighbor cells, leading to relatively long temporary disconnection of the transmission. In this paper, instead of the received signal strength, we introduce a stochastic connectivity estimation for selecting a proper target cell. The user connectivity to the BS is mainly related to the distance from the transmitter. Furthermore, stochastic factors such as shadowing and multipath fading influence the connectivity. If the received signal needs to be greater than $p_{0,\rm dB}$ for decoding data reliably, the connection probability can be obtained as follows [28]:

$$\begin{aligned} P_i^{\rm c}(j) &= Pr[p_{\rm t,dB} - \bar{L}_{0,\rm dB} - 10\log_{10} E[\chi^2] \\ &\quad - 10\xi \log_{10} D - X_{\sigma_{\rm s}} \geq p_{0,\rm dB}] \\ &= \frac{1}{2}(1 - {\rm erf}[(10\xi\log_{10} D + p_{0,\rm dB} - p_{\rm t,dB} \\ &\quad - \bar{L}_{0,\rm dB} - 10\log_{10} E[\chi^2])/\sqrt{2}\sigma_{\rm s}]), \end{aligned} \tag{11}$$

where $p_{\rm t,dB}$ is the transmission power, $\bar{L}_{0,\rm dB}$ is the average path loss at the reference distance, D is the distance from the BS, $10\log_{10} E[\chi^2]$ is the average multipath fading in dB, ξ is the path loss exponent, $X_{\sigma_{\rm s}}$ is shadowing, χ^2 is multipath fading, and ${\rm erf}[z]$ is the error function defined by $\int_0^z \frac{2}{\sqrt{\pi}} e^{-x^2} dx$.

Since the spectrum pool consists of multiple spectrum bands, the connectivity of spectrum pool i, $P_i^{\rm c}$, can be defined as the probability that at least one spectrum band provides the valid connection, which can be expressed as $1 - \prod_{j\in\mathcal{S}_i}(1 - P_i^{\rm c}(j))$ where $P_i^{\rm c}(j)$ is the connection probability of spectrum j in pool i. Besides the connectivity, spectrum utilization is also an important factor in determining the target cell. Thus, CR users select the target cell i^* with the highest weighted connectivity, $P_i^{\rm w}$, which can be obtained by considering both connectivity and spectrum utilization as follows:

$$P_i^{\rm w} = \left(1 - \prod_{j\in\mathcal{S}_i}(1 - P_i^{\rm c}(j))\right) \cdot \left(1 - \frac{N_i^{\rm b} + \rho N_i^{\rm e}}{\sum_{j\in\mathcal{S}_i} N_i^{\max}(j)}\right). \tag{12}$$

Note that the target cell should be selected among candidate cells that have the cell connectivity $P_i^{\rm c}$ supporting the minimum QoS for a mobile user.

6 User Mobility Management in Cognitive Radio Networks

6.1 Overview

The user mobility is another main reason to initiate handoff in CR networks, which happens at the boundary of either BA or EA.

- When CR users approach the boundary of EA, they check the feasibility of intercell/intrapool handoff (Type 5) first. Unlike the intercell/interpool handoff, CR users can measure the signal strength from other BS directly, which is exactly same as classical handoff schemes. If CR users cannot find a proper target cell for intercell/intrapool handoff, they need to perform the intercell/interpool handoff to find a cell having a different spectrum pool. This procedure is same as the cell selection scheme but does not require a preparation time (Type 3), which is explained in Section 5.3.
- When CR users approach the basic cell boundary, they need to determine whether they will stay in the EA of the current cell. For mobile users, a larger cell coverage is generally known to be much more advantageous since it reduces the number of handoffs [10]. However, in CR networks, the large cell coverage is not always desirable for mobile users. As the cell coverage becomes larger, the PU activity becomes higher since it is more highly probable to include multiple PU activity regions. The PU activity in EA results in a significantly long switching latency, as described in Section 4. In addition, since CR users in BA are allowed to have a higher priority in channel access, as presented in Section 5, cell overload also influences the use of extended spectrum band. As a result, CR networks need a sophisticated algorithm to select the best handoff type for mobile users at the boundary of BA.

Thus, in this section, we focus on the mobility management in the boundary of BA. When CR users become closer to the boundary, the BS initiates the handoff procedures and gather the neighbor cell information from a central network entity. Based on the information, the BS estimates the connectivity of the candidate cells and determines the handoff timing t^* and target cell i^* as follows:

$$[t^*, i^*] = \arg_{i\in\mathcal{C}, t>0} \left[P_{i_{\rm c}}^{\rm c}\left[d_{i_{\rm c}}^0 + v_{i_{\rm c}}^{\rm r} t\right] \leq \max_{i\in\mathcal{C}}\left[P_i^{\rm c}\left[d_i^0 + v_i^{\rm r} t\right]\right]\right], \tag{13}$$

where $P_i^{\rm c}$ is a connectivity of cell i, which is a function of the distance d_i^0 and the relevant velocity $v_i^{\rm r}$ to its BS. $\mathcal{C}$ is a set of candidate cells, $i_{\rm c}$ represents the current cell, and t is the moving time.

Once a target cell is determined, the BS determines the handoff type by considering the expected switching costs of

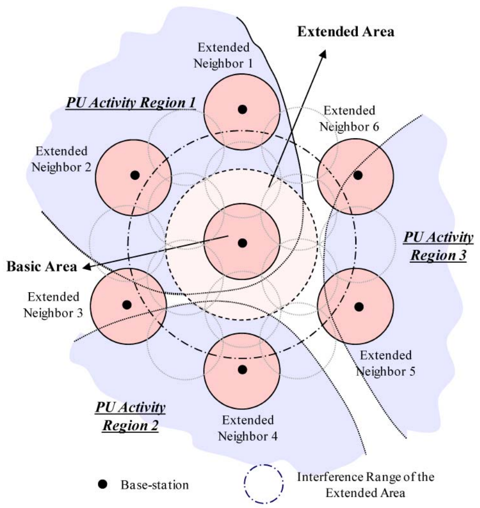


Fig. 4. The influence of primary user activities in the extended area.

both intracell/intrapool (Type 1) and intercell/interpool handoff schemes (Type 3) at the boundary of BA. The expected switching costs can be determined by estimating the probability of mobility events after the decision. After the decision, CR users may experience the unexpected intercell/interpool handoff due to the following reasons: 1) PU activity in EA, 2) capacity overload in EA, and 3) capacity overload in BA. In the following sections, first, we analyze these future events after the decision and accordingly propose an intelligent handoff decision scheme.

6.2 Primary User Activity in the Extended Area

If CR users determine to perform the intracell/intrapool handoff (Type 1) at the boundary of BA, they can stay in the current cell, which does not require a long switching latency for intercell/interpool handoff. However, in EA, CR users may experience the mobility events that cause intercell/interpool handoff (Type 4). One of those events is the PU activity. Since CR users in EA cannot find other available bands when they detect the PU activity, the intercell/interpool handoff is inevitable.

As shown in Fig. 4, more PU activity regions can be involved in determining spectrum availability in EA, which leads to a higher PU activity. Furthermore, the interference range of the extended spectrum is larger than its coverage, and hence is overlapped with the BAs of the extended neighbors. Thus, for the accurate detection, all extended neighbors need to be involved in detecting the PU activity with its own detection and false alarm probabilities. Assume that cooperative detection is performed according to the data fusion by OR rule [29], [30]. Then, a cooperative detection probability converges to 1 as the number of cells increases. Thus, the detection probability can be ignored when estimating the spectrum availability. On the contrary, the false alarm probability increases as the number of cells increases, which influences the spectrum availability significantly in EA. Even though a spectrum band is idle, it is determined to be unavailable if the false alarm is detected.

Thus, in order to avoid the intercell/interpool handoff, any PU activities or false alarms should not be detected in EA. Based on these observations, we derive the probability that no primary user can be detected during r sensing periods as follows:

$$P_i^{\mathrm{av}}(1) = \prod_{i' \in \mathcal{N}_i^{\mathrm{E}}} \left(1 - P_{i'}^{\mathrm{f}}\right) \cdot \prod_{k \in \mathcal{A}_i^{\mathrm{E}}(j)} e^{-\beta(j,k)\Delta t}, \tag{14}$$

$$P_i^{\mathrm{av}}(r) = \prod_{i' \in \mathcal{N}_i^{\mathrm{E}}} \left(1 - P_{i'}^{\mathrm{f}}\right)^r \cdot \prod_{k \in \mathcal{A}_i^{\mathrm{E}}(j)} e^{-\beta(j,k)r\Delta t} \cdot P_{\mathrm{E},i}^{\mathrm{under}}(r), \quad r = 2, 3, \ldots, R, \tag{15}$$

where $R = \lceil T_m/\Delta t \rceil$ where T_m is the expected time for user m to stay in EA. The first term represents the probability that all extended neighbors do not generate any false alarms during r sensing periods. This is based on the OR rule in the decision fusion, and will change if other cooperative decision criteria are used. Here, the sensing operation is assumed to be performed in every Δt sensing period. The second term denotes the probability that no PU activity appears in EA during r sensing slots. $P_{\mathrm{E},i}^{\mathrm{under}}(r)$ is the probability that the cell does not experience the capacity overload during r sensing slots, which is derived in the following section.

Then, the probability of intercell/interpool handoff due to the PU activity in EA can be obtained as follows:

$$P_{\mathrm{E},i}^{\mathrm{h_{pu}}} = \sum_{r=2}^{R} P_i^{\mathrm{av}}(r-1) \cdot \left(1 - P_i^{\mathrm{av}}(1)\right). \tag{16}$$

The first term in the summation represents the probability that the extended spectrum is available without cell overload during $r-1$ sensing slots, which is multiplied by the probability that the PU activity is detected at slot r.

6.3 Capacity Overload in the Extended Area

As explained in Section 5, when a current cell is overloaded, CR users in EA may need to perform an intercell/interpool handoff (Type 2). In this section, we derive a probability of cell overload. The PU activity in the extended spectrum leads to the intercell/interpool handoff regardless of the cell overload, which is already considered in Section 6.2. Thus, we assume that the cell overload results from PU activities only in basic spectrum bands, and the extended spectrum is considered to be idle in this case.

First, since each PU activity region in the spectrum has two states, busy and idle, we can model a transition matrix $\mathbf{X}(j,k)$ with following transition probabilities:

$$\begin{aligned} x_{1,1}(j,k) &= e^{-\beta(j,k)\Delta t}, \\ x_{1,2}(j,k) &= 1 - e^{-\beta(j,k)\Delta t}, \\ x_{2,1}(j,k) &= 1 - e^{-\alpha(j,k)\Delta t}, \\ x_{2,2}(j,k) &= e^{-\alpha(j,k)\Delta t}, \end{aligned} \tag{17}$$

where $x_{1,1}(j,k)$ and $x_{1,2}(j,k)$ are the transition probabilities from idle to idle and from idle to busy, respectively. $x_{2,1}(j,k)$ and $x_{2,2}(j,k)$ represent the transition probabilities from busy to idle and from busy to busy, respectively.

From this, the transition matrix after $r\Delta t$ can be obtained as $[\mathbf{X}(j,k)]^r$. Let $\mathbf{x_0}(j,k) \in \{(1,0),(0,1)\}$ be an initial vector to describe a current spectrum status where $(1,0)$ and $(0,1)$ denote that PU activity region k at spectrum j is currently idle and busy, respectively. Then, the idle probability of region k after $r\Delta t$, $P_i^{\text{idle}}(j,k,r\Delta)$ is the first element of the vector, $x_0[\mathbf{X}(j,k)]^r$, which can be obtained by (18) [31].

$$P_i^{\text{idle}}(j,k,r) = \begin{cases} \dfrac{x_{2,1}(j,k)}{x_{1,2}(j,k)+x_{2,1}(j,k)} + (1-x_{1,2}(j,k) & x_0=(1,0) \\ \quad -x_{2,1}(j,k))^r \cdot \dfrac{x_{1,2}(j,k)}{x_{1,2}(j,k)+x_{2,1}(j,k)}, & \\ \dfrac{x_{2,1}(j,k)}{x_{1,2}(j,k)+x_{2,1}(j,k)} - (1-x_{1,2}(j,k) & x_0=(0,1) \\ \quad -x_{2,1}(j,k))^r \cdot \dfrac{x_{2,1}(j,k)}{x_{1,2}(j,k)+x_{2,1}(j,k)}, & \end{cases} \tag{18}$$

Based on idle probabilities at each PU activity region, we derive the idle and busy probabilities of spectrum j. Assume that a current cell i has multiple PU activity regions in spectrum j. Then, it can use that spectrum only when all of these regions are idle, and hence the idle and busy probabilities are expressed as follows:

$$\begin{aligned} P_i^{\text{idle}}(j,r) &= \prod_{k\in\mathcal{A}_i^{\text{B}}(j)} P_i^{\text{idle}}(j,k,r), \\ P_i^{\text{busy}}(j,r) &= 1-P_i^{\text{idle}}(j,r), \end{aligned} \tag{19}$$

where $\mathcal{A}_i^{\text{B}}(j)$ is the set of PU activity regions of BA in spectrum j at cell i.

Based on both probabilities of each spectrum in the pool, we derive the overload probability of cell i as follows: assume that $\tilde{S}_i$ is a set of all spectrum bands assigned to the current cell i, which are used by either primary or CR users. Since the extended spectrum is not considered as explained earlier, the cell has $2^{|\tilde{S}_i|-1}$ states according to spectrum availability. Among them, some states cannot satisfy capacity requirements to support current users in the cell, resulting in cell overload. These states can be obtained as follows:

$$\mathcal{L}_{\text{E}} = \left\{ \arg_n \left[\sum_{j\in\mathcal{I}_n} N_i^{\max}(j) < N_i^{\text{b}} + \rho N_i^{\text{e}} \right], \text{ for } \forall n \right\}, \tag{20}$$

where $\mathcal{I}_n$ is a set of idle spectrum bands at state n $(n=1,\ldots,2^{|\tilde{S}_i|-1})$.

To resolve cell overload at each state in $\mathcal{L}_{\text{E}}$, the cell needs to obtain additional channels by switching some CR users to neighbor cells. The following is the probability that a mobile user c in EA is switched to the neighbor cell due to cell overload at state $n \in \mathcal{L}_{\text{E}}$:

$$\begin{aligned} u_{\text{E}}^{\text{h}}(r,n) = &\frac{\min\left[N_i^{\text{b}} + \rho N_i^{\text{e}} - B_i^{\max}(n), \rho N_i^{\text{e}}\right]}{\rho N_i^{\text{e}}} \\ &\cdot \frac{K_{\text{c}}(d_{\text{c}}/v_{\text{c}} - r\cdot\Delta t)}{\sum_{m\in\mathcal{U}_i^{\text{e}}} K_m d_m/v_m}, \end{aligned} \tag{21}$$

where $B_i^{\max}(n)$ is the number of available channels in cell i at state n, which can be obtained as $\sum_{j\in\mathcal{I}_n} N_i^{\max}(j)$. $N_i^{\text{b}} + \rho N_i^{\text{e}} - B_i^{\max}(n)$ represents the number of channels requiring the intercell/interpool handoff to prevent cell overload. K_{c} and $d_{\text{c}}/v_{\text{c}}$ are the QoS requirement of the current mobile user, and its expected staying time in EA, respectively. The first term in (27) represents the ratio of the overloaded capacity in EA. This ratio is multiplied by a weighted coefficient to consider the spectrum mobility function where the cell selects users having a higher selection metric $K_m \cdot d_m/v_m$, as presented in Section 5.2. As shown in the second term of (27), the weighted coefficient is defined as the ratio of user c's selection metric to the sum of those of all users in EA, and decreases as the user approaches the boundary of EA.

Let $\mathcal{I}_n$ be a set of idle spectrum bands at state $n \in \mathcal{L}_E$. Then, the probability of cell overload at rth sensing slot can be obtained as follows:

$$P_{\text{E},i}^{\text{over}}(r) = \sum_{n\in\mathcal{L}_{\text{E}}} \left[\prod_{j\in\mathcal{I}_n} P_i^{\text{idle}}(j,r) \prod_{j\notin\mathcal{I}_n} P_i^{\text{busy}}(j,r) \cdot u_{\text{E}}^{\text{h}}(r,n) \right]. \tag{22}$$

Accordingly, the probability that the cell is not overloaded $P_{\text{E},i}^{\text{under}}(r)$ is derived as $1-P_{\text{E},i}^{\text{over}}(r)$.

To avoid intercell/interpool handoff in EA, the cell should not experience any cell overload as well as any PU activity during their staying time. Thus, its probability during $r\Delta t$ can be expressed as follows:

$$P_{\text{E},i}^{\text{h}_{\text{no}}}(r) = \left(P_i^{\text{av}}(r) - P_{\text{E},i}^{\text{h}_{\text{over}}}(r)\right), \; r=1,2,\ldots,R. \tag{23}$$

Here, $P_{\text{E},i}^{\text{h}_{\text{over}}}(r)$ is the probability of intercell/interpool handoff in EA caused by cell overload during $r\Delta t$, which is expressed as follows:

$$P_{\text{E},i}^{\text{h}_{\text{over}}}(1) = P_i^{\text{av}}(1) \cdot P_{\text{E},i}^{\text{over}}(1), \tag{24}$$

$$\begin{aligned} P_{\text{E},i}^{\text{h}_{\text{over}}}(r) = &P_i^{\text{av}}(1) \cdot P_{\text{E},i}^{\text{h}_{\text{no}}}(r-1) P_{\text{E},i}^{\text{over}}(r) \\ &+ P_{\text{E},i}^{\text{h}_{\text{over}}}(r-1), \; r=2,3,\ldots,R. \end{aligned} \tag{25}$$

The first term in (25) represents the probability that intercell/interpool handoff due to cell overload occurs in sensing slot r, which is obtained by multiplying the probability that no PU activity is detected in slot r, the probability that no intercell/interpool handoff occurs during $r-1$ slots, and the cell overload probability in slot r.

6.4 Capacity Overload in the Basic Area

If the BS determines to perform the intercell/interpool handoff (Type 3) at the boundary of BA, mobile CR users may experience the capacity overload in the BA of target cell, which causes intercell/interpool handoff. This cell overload probability can be determined with the similar procedures used in deriving $P_{\text{E}}^{\text{over}}$ in Section 6.3.

First, the spectrum states detecting cell overload in BA can be derived as follows:

$$\mathcal{L}_{\text{B}} = \left\{ \arg_n \left[\sum_{j\in\mathcal{I}_n} N_i^{\max}(j) < N_i^{\text{b}} \right], \quad \text{for } \forall n \right\}. \tag{26}$$

Based on overload states $n \in \mathcal{L}_{\text{B}}$, we derive a probability of intercell/interpool handoff in BA to resolve cell overload as follows:

$$u_{\mathrm{B}}^{\mathrm{ho}}(r,n) = \frac{\min\left[N_i^{\mathrm{b}} - N_i^{\max}(n), N_i^{\mathrm{b}}\right]}{N_i^{\mathrm{b}}} \cdot K_{\mathrm{c}} \cdot \left(1 - \frac{d_{\mathrm{c}}/v_{\mathrm{c}} - r \cdot \Delta t}{\sum_{m \in \mathcal{U}_i^e} K_m d_m / v_m}\right). \quad (27)$$

The first term is the ratio of the number of channels in BA required to resolve cell overload to total channels of current users in BA. In the case of cell overload in BA, users having a shorter expected staying time are selected for intercell/interpool handoff, as explained in Section 5.2. Thus, the weighed coefficient in the second term of (27) is expressed as one minus the ratio of user c's selection metric to the sum of those of all users in BA, which is different from (27).

The probability of cell overload in BA, $P_{\mathrm{B},i}^{\mathrm{over}}(r)$, can be obtained by replacing $\mathcal{L}_{\mathrm{E}}$ and $u_{\mathrm{E}}^{\mathrm{h}}(n)$ in (22) with $\mathcal{L}_{\mathrm{B}}$ and $u_{\mathrm{B}}^{\mathrm{h}}(n)$, respectively. Accordingly, the probability that the CR users in BA perform intercell/interpool handoff during r slots, $P_{\mathrm{B}}^{\mathrm{h_{over}}}(r)$, is estimated as follows:

$$P_{\mathrm{B},i}^{\mathrm{h_{over}}}(1) = P_{\mathrm{B},i}^{\mathrm{over}}(1), \quad (28)$$

$$P_{\mathrm{B},i}^{\mathrm{h_{over}}}(r) = \left(1 - P_{\mathrm{B},i}^{\mathrm{h_{over}}}(r-1)\right) \cdot P_{\mathrm{B},i}^{\mathrm{over}}(r) + P_{\mathrm{B},i}^{\mathrm{h_{over}}}(r-1),\ r = 2, 3, \ldots, R. \quad (29)$$

Unlike (25), we consider all spectrum bands including the extended spectrum in this case. Thus, we do not need to consider the probability of the spectrum availability in EA, $P_i^{\mathrm{av}}(r)$.

6.5 Switching Cost

According to the probability on future mobility events, we estimate the switching cost of two possible options in the boundary of BA. First, when CR users stay in the current cell by performing intracell/intrapool handoff to EA, the expected switching cost T_{EA} can be obtained as follows:

$$T_{\mathrm{EA}} = D_1 + P_{\mathrm{E},i}^{\mathrm{h_{over}}}(R) \cdot D_2 + P_{\mathrm{E},i}^{\mathrm{h_{pu}}} \cdot D_4 + \left(1 - P_{\mathrm{E},i}^{\mathrm{h_{over}}}(R) - P_{\mathrm{E},i}^{\mathrm{h_{pu}}}\right) \cdot D_5. \quad (30)$$

The total delay includes the intracell/intrapool handoff when the CR user switches to the extended spectrum, intercell/interpool handoffs due to the cell overload and and PU activity, and intercell/intrapool handoff when it is successfully handed over to the extended neighbors.

Second, when CR users move to the neighbor cell by performing intercell/interpool handoff, the expected switching cost can be expressed as the sum of the instant switching delay and the expected switching delay due to the overload in that neighbor cell as follows:

$$T_{\mathrm{BA}} = D_3 + D_1 \frac{T_m}{\overline{T}_{\mathrm{off},i} + D_1} + P_{\mathrm{B},i}^{\mathrm{h_{over}}}(R) \cdot D_2. \quad (31)$$

The latency in this case includes the intercell/interpool handoff to a new target cell , intracell/intrapool handoff in the target cell, and intercell/interpool handoff due to cell overload. Here, the average number of intracell/intrapool handoff is obtained as $T_m/(\overline{T}_{\mathrm{off},i} + D_1)$. $\overline{T}_{\mathrm{off},i}$ is the average idle period of the spectrum in cell i, which is expressed as the average of $1/\sum_{k \in \mathcal{A}_i^{\mathrm{B}}(j)} \beta(j,k)$ over all spectrum j.

TABLE 2
Handoff Delay Components Used in Simulations

Components	d_{prep}	d_{recfg}	d_{lis}	$d_{\mathrm{syn}}^{\mathrm{sen}}$	d_{sen}	d_{dec}	$d_{\mathrm{syn}}^{\mathrm{tx}}$
Delay (sec)	0.1	0.3	0.1	0.025	0.025	0.025	0.025

Based on the analysis above, the BS determines the handoff type with the lower expected spectrum cost.

7 Performance Evaluation

7.1 Simulation Setup

In order to evaluate the performance of the proposed mobility management framework, we implement a network simulator to support the network topology consisting of multiple cells in $10\,\mathrm{km} \times 10\,\mathrm{km}$ area. Here, we assume 59 cells that have different channel utilization. The transmission range of each cell is set to 750 m. The interference range is set to twice larger than the transmission range. The transmission range of the extended spectrum is also twice larger than that of basic spectrum. Furthermore, we consider four spectrum pools, each of which consists of 10 spectrum bands. The basic and extended spectrum bands can support 10 and 40 channels for users in BA, respectively (i.e., ρ is set to 4). Furthermore, each band has three to five PU activity regions, which have different PU activities, $\alpha(j,k)$ and $\beta(j,k)$ uniformly distributed in [0.01, 0.05]. The sensing interval Δt is 0.1 sec. The BSs are assumed to generate a false alarm every 2 hours on average when they sense the availability of each spectrum.

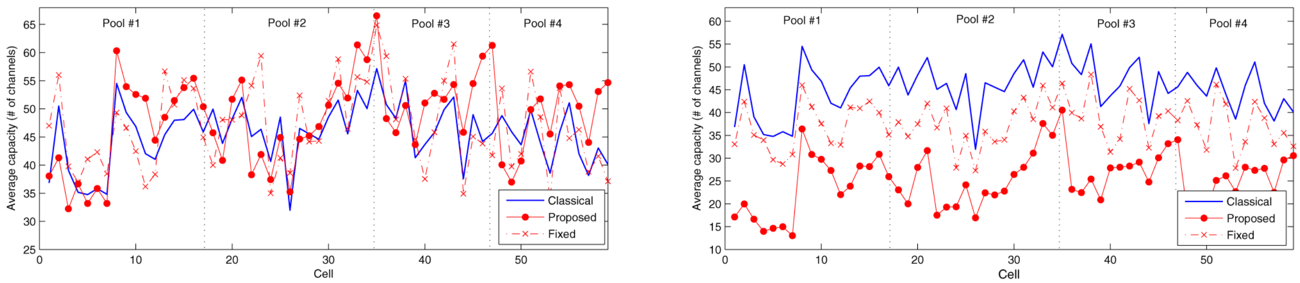
Furthermore, based on the delay components in Table 2, the handoff delays defined in Section 4, D_1, D_2, D_3, D_4, and D_5 are set to 0.2, 0.5, 0.4, 1.2, and 0.1 sec, respectively. An operational delay for intercell resource allocation, T^{inter}, is assumed to be 5 sec.

In a power attenuation model, a channel gain is set to -31.54 dB, a reference distance is 1 meter, and a path loss coefficient is 3.5. Shadow fading σ_s and multipath fading of spectrum bands in each cell are randomly distributed in [3, 7] and [2, 4] dB, respectively. The BS uses $-56.21\,\mathrm{dBm/Hz}$ transmission power on average for the basic spectrum and $-47.18\,\mathrm{dBm/Hz}$ for the extended spectrum. Noise power in the receiver is $-174\,\mathrm{dBm/Hz}$. The minimum decodable SNR is set to 0 dB.

To describe user mobility, we consider a Gauss-Markov mobility model proposed in [32], where the memory level parameter is set to 0.5, and asymptotic standard deviations are 1.2, 6, 12, and 20 m/sec for average velocities 6, 30, 60, and 100 km/h, respectively. Note that any mobility model can be applied to the proposed method, but according to its accuracy, the performance of the proposed method changes significantly.

7.2 Performance of Intercell Resource Allocation for Extended Spectrum Bands

In this simulation, we evaluate the performance of the proposed intercell resource allocation by comparing with the following methods:



(a) (b)

Fig. 5. Average channel availability: (a) best case and (b) worst case.

- *Classical handoff scheme.* This scheme supports only the basic spectrum bands, and does not have an extended spectrum. Thus, each cell is able to access all available spectrum bands in the pool without influence on its extended neighbors.
- *Highest capacity preferred scheme.* The BS selects the extended spectrum so as to maximize the total number of available channels in the network. A decision principle of this scheme is similar to (3), but does not consider a reliability metric $R_i(j)$.
- *Highest availability preferred scheme.* The spectrum with the highest idle probability is selected for the extended spectrum, i.e., $G_i(j)$ is set to $\prod_{k\in\mathcal{A}_i^{\mathrm{E}}(j)} P_i^{\mathrm{off}}(j,k)$ in (4).
- *Fixed allocation.* Each cell is assigned to the predetermined extended spectrum bands based on the proposed method in (3), but does not change them regardless of time-varying spectrum availability.

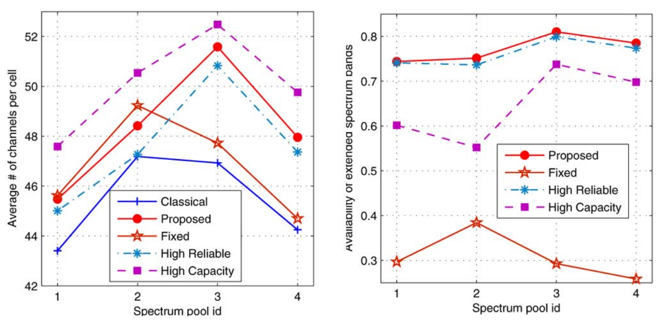
In Fig. 5, we investigate total spectrum availability, i.e., the total network capacity of each scheme. If the cell has the extended spectrum, total network capacity is dependent on the location of users. Figs. 5a and 5b show the best case (i.e., all users using the extended spectrum are located on BA), and the worst case (i.e., all users in the extended spectrum are located on EA), respectively. In the best case, both proposed and fixed methods show slightly higher capacity than the classical approach since the extended spectrum supports more channels for users in BA although it restricts the use of that spectrum in its extended neighbors. On the contrary, in the worst case, the classical method has much more available channels because the use of the extended spectrum in both proposed and fixed methods reduces the channel utilization in extended neighbor cells while users in EA require more channel resource for the same quality of service as users in BA. In this case, since the proposed method has a higher utilization of the extended spectrum, it shows the lowest number of available channels.

In Fig. 6, we compare the proposed method with other decision principles in terms of total capacity (best case) and the availability of the extended spectrum. The availability of the extended spectrum is defined as the ratio of the time that the extended spectrum is valid for the cell to total simulation time. The highest capacity preferred method shows the highest total channel availability by reducing the effect on the rest of networks, but has trouble with finding more reliable extended spectrum. The highest availability preferred scheme shows lower capacity since it causes an adverse influence on neighbor cells. In addition, since it only focuses on the overall idle probability of the spectrum without the consideration of intercell operational delay, it may choose the spectrum requiring more frequent switching, leading to lower reliability in the extended spectrum than the proposed method. On the contrary, the proposed method shows the highest availability of the extended spectrum while maintaining higher capacity compared to the highest availability preferred and fixed methods by jointly considering a capacity gain and reliability in the extended spectrum.

In summary, the use of the extended spectrum leads to lower network capacity compared with the classical methods but higher availability in the extended spectrum. However, it improves mobility performance in CR cellular network, and hence allows the proposed method to achieve higher actual total capacity by reducing the adverse effects of dynamic network environments, which is shown in the subsequent simulations.

7.3 Performance of Spectrum and User Mobility Management Schemes

In this simulation, we investigate transmission statistics in mobile users under different network environments to evaluate the performance of both spectrum and mobility management schemes. To this end, we perform 20 1-hour simulations for each case and obtain average values. Here, we analyze the performance of mobility management in terms of three factors: user QoS requirement (i.e., the number of channels required for a current communication), current network load (i.e., the number of channels currently occupied by other users), and the velocity of



(a) (b)

Fig. 6. Intercell resource allocation: (a) total available channels and (b) availability in extended spectrum bands.

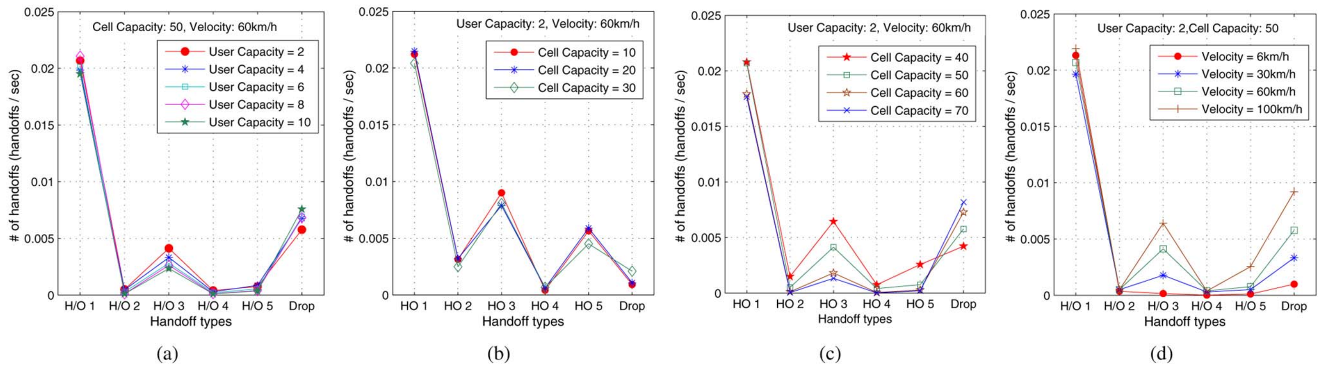


Fig. 7. Handoff types: (a) user capacity, (b) lower cell occupancy, (c) higher cell occupancy, and (d) user velocity.

mobile users. In the simulation, we compare the proposed method with the classical handoff scheme and two other methods as follows:

- *BA-preferred.* This scheme has the extended spectrum, the use of which aims at increase in the number of channels in BA. At the cell boundary, a mobile user switches to its neighbor cell if a valid target cell exist. Otherwise, it moves to EA.
- *EA-preferred.* This scheme mainly focuses on enhancing mobility. Thus, a mobile user at the cell boundary tries to stay in the EA of the current cell if the extended spectrum is available.

Fig. 7 shows the number of different mobility events in the proposed method. As the user QoS requirement increases, the number of each handoff type decreases since it reduces the probability to find enough resource in BA and EA (Fig. 7a). Figs. 7b and 7c show the changes in handoff types according to network load. If the network is underloaded, the cell capacity does not influence the number of each handoff type significantly. On the contrary, in a highly loaded network, as the cell capacity increases, the number of all types of handoff decreases and conversely a drop rate increases because increase in the cell overload probability reduces transmission opportunity, which is explained in Fig. 7c. Furthermore, the underloaded case can maintain a higher availability of the extended spectrum due to a lower cell overload probability, and hence shows more Type 5 handoffs than the highly loaded case. If user velocity increases, Type 5 handoff to EA increases to reduce the abrupt quality degradation due to the frequent intercell/interpool handoff. In all cases, the proposed method sustains the lowest number of the worst handoff (Type 4) by intelligently choosing proper handoff types based on the expected switching delay.

One of the most important statistics in mobility management is a call drop probability. The call drop occurs when a mobile user cannot find any available spectrum in both serving and target cells. Here, we do not consider a call blocking probability. Fig. 8 shows simulation results on a drop rate under different situations. To show the reliability of simulation results, we indicate 95 percent confidence interval on the graph of the proposed method. In Fig. 8, the proposed method shows better performance in the drop rate than classical and other handoff methods. As shown in Fig. 8a, although the user QoS requirement increases, the proposed method maintains a certain level of drop rate through spectrum mobility management. If the network load increases, a drop rate becomes higher due to the lack of available spectrum resource, but is still lower than other methods by selecting the handoff type adaptively to cell conditions. Furthermore, the proposed method allows mobile users to adaptively switch to either BA or EA while reducing the number of intercell/interpool handoff. As a result, the proposed method sustains a lower drop rate although a mobile user traverses across more cell boundaries with a higher velocity, as shown in Fig. 8c.

Fig. 9 shows the link efficiency, which is defined as a real transmission time over an entire simulation time. In this simulation, the classical method shows a lower link efficiency over all cases due to quality degradation caused by frequent intercell/interpool handoffs. Furthermore,

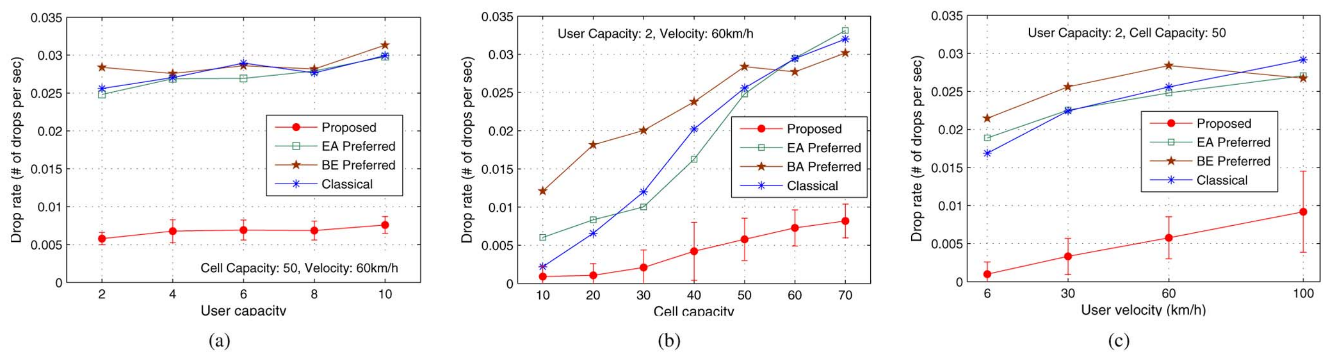
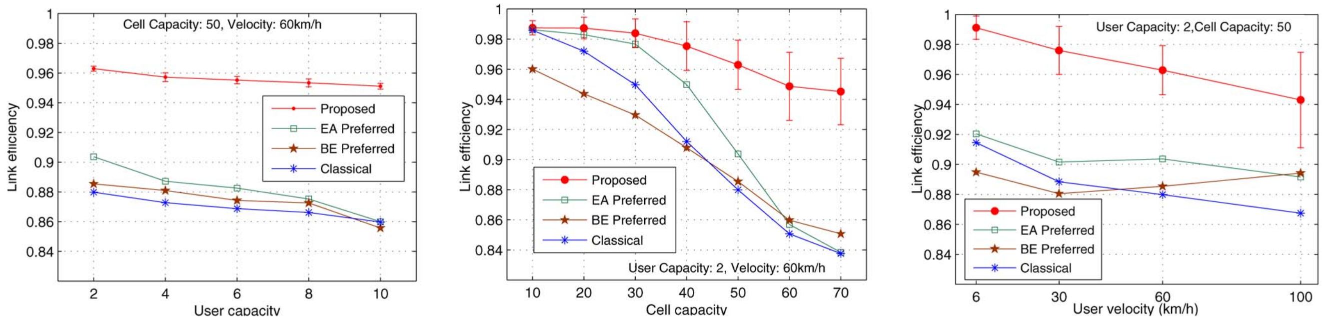


Fig. 8. Drop rate: (a) user capacity, (b) cell occupancy, and (c) user velocity.



(a) (b) (c)

Fig. 9. Link efficiency: (a) user capacity, (b) cell occupancy, and (c) user velocity.

when the current cell is overloaded, some mobile users cannot use spectrum resources until spectrum availability changes or they move into a new target cell, which also reduces the link efficiency. Both EA and BA preferred methods also show lower link efficiencies since they do not consider the expected switching latency caused by future mobility events. On the contrary, the proposed method shows a higher link efficiency by intelligently determining the handoff type to reduce the latency as well as the drop rate.

From these simulations, we can see that the proposed method achieves more actual transmission opportunity as well as less quality degradation during the transmission regardless of user and network conditions, although it shows lower network capacity theoretically due to the use of extended spectrum bands.

8 Conclusions

In this paper, we present a spectrum-aware mobility management scheme for CR cellular networks. Available spectrum bands in CR cellular networks vary over time and space and are distributed discontinuously over a wide frequency range. First, we propose the spectrum pool-based network architecture, which mitigates the heterogeneous spectrum availability. Based on this architecture, a unified mobility management framework is defined so as to support diverse mobility events in CR networks, consisting of intercell resource allocation, and spectrum and user mobility management functions. Through intercell resource allocation, each cell determines its spectrum configuration to improve mobility as well as total capacity. For the PU activity, spectrum mobility management is developed where the network determines a proper spectrum band and target cell according to both current spectrum utilization and stochastic connectivity model. In user mobility management, the switching cost-based handoff decision mechanism is proposed so as to minimize quality degradation caused by user mobility. Simulation results show that the proposed method provides maximum cell capacity while providing minimum quality degradation in mobile users.

Acknowledgments

This material is based upon work supported by the US National Science Foundation under Grant No. CNS-07251580.

References

[1] Federal Communications Comission, “Spectrum Policy Task Force Report,” ET Docket No. 02-135, Nov. 2002.

[2] I.F. Akyildiz, W.Y. Lee, M.C. Vuran, and S. Mohanty, “Next Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey,” *Computer Networks*, vol. 50, pp. 2127-2159, Sept. 2006.

[3] *IEEE P802.22/D0.3.8.1, IEEE 802.22 WG, Draft Standard for Wireless Regional Area Networks Part 22: Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE, Sept. 2007.

[4] M.M. Buddhikot, “Cognitive Radio, DSA and Self-X: Towards Next Transformation in Cellular Networks,” *Proc. IEEE Symp. New Frontiers in Dynamic Spectrum (DySPAN ’10)*, Apr. 2010.

[5] J. Sachs, I. Maric, and A. Goldsmith, “Cognitive Cellular Systems within the TV Spectrum,” *Proc. IEEE Symp. New Frontiers in Dynamic Spectrum (DySPAN ’10)*, Apr. 2010.

[6] L. Yang, L. Cao, and H. Zheng, “Proactive Channel Access in Dynamic Spectrum Network,” *Proc. Int’l ICST Conf. Cognitive Radio Oriented Wireless Networks (CROWNCOM ’07)*, July 2007.

[7] X. Zhu, L. Shen, and T.P. Yum, “Analysis of Cognitive Radio Spectrum Access with Optimal Channel Reservation,” *IEEE Comm. Letters*, vol. 11, no. 4, pp. 304-306, Apr. 2007.

[8] H. Kim and K.G. Shin, “Fast Discovery of Spectrum Opportunities in Cognitive Radio Networks,” *Proc. IEEE Symp. New Frontiers in Dynamic Spectrum Access Networks (DySPAN ’08)*, Oct. 2008.

[9] I.F. Akyildiz, J. McNair, J.S.M. Ho, H. Uzunalioglu, and W. Wang, “Mobility Management in Next-Generation Wireless Systems,” *Proc. IEEE*, vol. 87, no. 8, pp. 1347-1384, Aug. 1999.

[10] P.A. Ramsdale and W.B. Harrold, “Techniques for Cellular Networks Incorporating Microcells,” *Proc. IEEE Int’l Symp. Personal, Indoor and Mobile Radio Comm. (PIMRC ’92)*, pp. 169-173, Oct. 1992.

[11] S.V. Hanly, “An Algorithm for Combined Cell-Site Selection and Power Control to Maximize Cellular Spread Spectrum Capacity,” *IEEE J. Selected Areas in Comm.*, vol. 13, no. 7, pp. 1332-1340, Sept. 1995.

[12] A. Sang, X. Wang, M. Madihian, and R.D. Gitlin, “Coordinated Load Balancing, Handoff/Cell-Site Selection, and Scheduling in Multi-Cell Packet Data Systems,” *Proc. ACM MobiCom*, pp. 302-314, Sept. 2004.

[13] D. Amzallag, R. Bar-Yehuda, D. Raz, and G. Scalosub, “Cell Selection in 4G Cellular Networks,” *Proc. IEEE INFOCOM*, pp. 1373-1381, Apr. 2008.

[14] Federal Communications Comission, “In the Matter of Unlicensed Operation in the TV Broadcast Bands: Second Report and Order and Memorandum Opinion and Order,” FCC 08-260, Nov. 2008.

[15] 3GPP TSG-RAN, “Requirements for Further Advancements for Evolved Universal Terrestrial Radio Access (E-UTRA),” 3GPP TR 36.913 V9.0.0, Dec. 2008.

[16] M.M. Buddhikot, I. Kennedy, F. Mullany, and H. Viswanathan, “Ultrabroadband Femtocells via Opportunistic Reuse of Multi-Operator and Multi-Service Spectrum,” *Bell Labs Technical J.*, Special Issue on 4G Networks, vol. 13, pp. 129-143, Feb. 2009.

[17] 3GPP TSG RAN WG4, “RF Requirements for Multicarrier and Multi-RAT BS (Release 9),” 3GPP TR 37.900 V1.0.0, Sept. 2009.

[18] R. Tandra, S.M. Mishra, and A. Sahai, "What Is a Spectrum Hole and What Does It Take to Recognize One?" *Proc. IEEE*, vol. 97, no. 5, pp. 824-848, May 2009.

[19] K. Sriram and W. Whitt, "Characterizing Superposition Arrival Processes in Packet Multiplexers for Voice and Data," *IEEE J. Selected Areas in Comm.*, vol. 4, no. 6, pp. 833-846, Sept. 1986.

[20] C. Chou, S. Shankar, H. Kim, and K.G. Shin, "What and How Much to Gain by Spectrum Agility?" *IEEE J. Selected Areas in Comm.*, vol. 25, no. 3, pp. 576-588, Apr. 2007.

[21] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentrallized Cognitive Mac Opportunistic Spectrum Access in Ad Hoc Networks: A Pomdp Framework," *IEEE J. Selected Areas in Comm.*, vol. 25, no. 3, pp. 589-600, Apr. 2007.

[22] T.A. Weiss and F.K. Jondral, "Spectrum Pooling: An Innovative Strategy for the Enhancement of Spectrum Efficiency," *IEEE Comm. Magazine*, vol. 42, no. 3, pp. 8-14, Mar. 2004.

[23] D. Cabric, S. Mishra, D. Willkomm, R. Brodersen, and A. Wolisz, "A Cognitive Radio Approach for Usage of Virtual Unlicensed Spectrum," *Proc. 14th IST Mobile and Wireless Comm. Summit*, June 2005.

[24] L. Cao and H. Zheng, "Distributed Spectrum Allocation via Local Bargaining," *Proc. IEEE Sensor and Ad Hoc Comm. and Networks (SECON '05)*, pp. 475-486, Sept. 2005.

[25] W.Y. Lee and I.F. Akyildiz, "Joint Spectrum and Power Allocation for Inter-Cell Spectrum Sharing in Cognitive Radio Networks," *Proc. IEEE Symp. New Frontiers in Dynamic Spectrum Access Networks (DySPAN '08)*, Oct. 2008.

[26] C. Peng, H. Zheng, and B.Y. Zhao, "Utilization and Fairness in Spectrum Assignment for Opportunistic Spectrum Access," *ACM Mobile Networks and Applications*, vol. 11, pp. 555-576, 2006.

[27] W.Y. Lee and I.F. Akyildiz, "A Spectrum Decision Framework for Cognitive Radio Networks," *IEEE Trans. Mobile Computing*, vol. 10, no. 2, pp. 161-174, Feb. 2011.

[28] W. Wang and M. Zhao, "Joint Effects of Radio Channels and Node Mobility on Link Dynamics in Wireless Networks," *Proc. IEEE INFOCOM*, pp. 933-941, Apr. 2008.

[29] Y.C. Liang, Y. Zeng, E. Peh, and A.T. Hoang, "Sensing-Throughput Tradeoff for Cognitive Radio Networks," *IEEE Trans. Wireless Comm.*, vol. 7, no. 4, pp. 1326-1337, Apr. 2008.

[30] S.M. Mishra, A. Sahai, and R.W. Brodersen, "Cooperative Sensing among Cognitive Radios," *Proc. IEEE Int'l Conf. Comm. (ICC '06)*, vol. 4, pp. 1658-1663, June 2006.

[31] D. Kannan, *An Introduction to Stochastic Processes.* North Holland, 1979.

[32] D. Liang and Z.J. Haas, "Predictive Distance-Based Mobility Management for PCS Networks," *Proc. IEEE INFOCOM*, vol. 3, pp. 1377-1384, Mar. 1999.

Won-Yeol Lee received the BS and MS degrees from the Department of Electronic Engineering, Yonsei University, Seoul, Korea, in 1997 and 1999, respectively. He received the PhD degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, in 2009 under the guidance of Prof. Ian F. Akyildiz. From 1999 to 2004, he was the senior research engineer of the Network R&D Center and Wireless Multimedia Service Development Division at LG Telecom, Seoul, Korea. Currently, he is the deputy director of the Next Mobile Technology Group, Network R&D Laboratory, Korea Telecom (KT), Seoul, Korea. His current research interests include cognitive radio networks, next generation wireless systems, and wireless sensor networks. He received the 2008 Researcher of the Year Award from the Broadband Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology. He is a student member of the IEEE.

Ian F. Akyildiz is the Ken Byers distinguished chair professor with the School of Electrical and Computer Engineering, Georgia Institute of Technology. Since June 2008, he has been an honorary professor with the School of Electrical Engineering at the Universitat Politecnico de Catalunya, Barcelona, Spain. Also since March 2009, he has been an honorary professor with the Department of Electrical, Electronic and Computer Engineering at the University of Pretoria, South Africa. He is the editor-in-chief of the *Computer Networks* (COMNET) journal as well as the founding editor-in-chief of the *Ad Hoc Networks* journal and the *Physical Communication* journal, all with Elsevier. His current research interests are in cognitive radio networks, wireless sensor networks, and nanocommunication networks. He has received numerous awards, including the 1997 IEEE Leonard G. Abraham Prize Award (IEEE Communications Society) for his paper entitled "Multimedia Group Synchronization Protocols for Integrated Services Architectures" published in the *IEEE JSAC* in January 1996; the 2002 IEEE Harry M. Goode Memorial Award (IEEE Computer Society) with the citation "for significant and pioneering contributions to advanced architectures and protocols for wireless and satellite networking"; the 2003 IEEE Best Tutorial Award (IEEE Communication Society) for his paper entitled "A Survey on Sensor Networks," published in *IEEE Communications* magazine in August 2002; the 2003 ACM Sigmobile Outstanding Contribution Award with the citation "for pioneering contributions in the area of mobility and resource management for wireless communication networks"; the 2004 Georgia Tech Faculty Research Author Award for his "outstanding record of publications of papers between 1999 and 2003; the 2005 Distinguished Faculty Achievement Award from School of Electrical and Computer Engineering, Georgia Tech; the 2009 Georgia Tech Outstanding Doctoral Thesis Advisor Award for "his 20+ years of service and dedication to Georgia Tech and producing outstanding PhD students"; and the 2009 Electrical and Computer Engineering Distinguished Mentor Award from the School of Electrical and Computer Engineering, Georgia Tech. He has been a fellow of the ACM since 1996. He is a fellow of the IEEE.